

Ethical challenges and innovations in the dissemination of genomic data: the experience of the PERSPECTIVE project

Emmanuelle Lévesque¹
Bartha Maria Knoppers¹
Jacques Simard²

¹Department of Human Genetics,
Centre for Genomics and Policy,
McGill University, Montréal,

²Genomics Centre, CHU de Québec
Research Center, Department of
Molecular Medicine, Laval University,
Québec City, QC, Canada

Abstract: The importance of making genomic data available for future research is now widely recognized among the scientific community and policymakers. In this era of shared responsibility for data dissemination, improved patient care through research depends on the development of powerful and secure data-sharing systems. As part of the concerted effort to share research resources, the project entitled Personalized Risk Stratification for Prevention and Early Detection of Breast Cancer (PERSPECTIVE) makes effective data sharing through the development of a data-sharing framework, one of its goals. The secondary uses of data from PERSPECTIVE for future research promise to enhance our knowledge of breast cancer etiologies without duplicating data-gathering efforts. Despite its benefit for research, we recognize the ethical challenges of data sharing on the local, national, and international levels. The effective management of ethical approvals for projects spanning across jurisdictions, the return of results to research participants, and research incentives and recognition for data production, are but a few pressing issues that need to be properly addressed. We discuss how we managed these issues and suggest how ongoing innovations might help to facilitate data sharing in future genomic research projects.

Keywords: data sharing, research ethics, cancer

Introduction

Recently, *Science* announced it will promote selected data repositories, tag their datasets, and integrate such data-tagging schemas in its publications.¹ The editor-in-chief asked readers to help identify repositories that are “well managed, have long-term support and are responsive to community needs”.¹ We are entering the era of shared responsibility for data dissemination. Research participants, funding agencies, data producers, data analysts, and now editors consider themselves part of this global data-sharing initiative. Opportunities to improve cancer patient care through genomic research, for example, depend on the development of powerful and secure data-sharing systems.²

Our international genomic research project on breast cancer, entitled Personalized Risk Stratification for Prevention and Early Detection of Breast Cancer (PERSPECTIVE), began in 2013. It builds large datasets benefiting from the support of its international research partners. Establishing data repository ethical frameworks for effective international data sharing over time is a challenge.

PERSPECTIVE is developing and implementing a personalized risk-stratification approach in order to prevent and detect breast cancer at its earliest stages. This approach seeks to classify women into risk categories using both genomic and phenotypic data. Genomic data will come from a genomic profiling test that

Correspondence: Emmanuelle Lévesque
Department of Human Genetics,
Centre for Genomics and Policy, McGill
University, 740, Dr Penfield Avenue,
Suite 5202, Montréal, QC, Canada,
H3A 1A4
Tel +1 514 398 8279
Fax +1 514 398 8954
Email emmanuelle.levesque@mcgill.ca

detects common and low-frequency genetic variants. Phenotypic data include various personal risk factors, such as breast density, age, familial history, hormone replacement therapy use, etc. As part of personalized medicine, the goal of this stratification is to offer women access to personalized, risk-adapted screening and prevention options. For instance, women at increased risk could be offered earlier and/or more screening than women at population risk.

Significant amounts of genomic data, and their associated phenotypic data, were collected and produced in order to develop such genomic screening tools. In fact, more than 150,000 samples and associated data generated from 90 studies³ around the world are being analyzed in PERSPECTIVE. These analyses will allow the development of a genomic profiling test to inform a risk prediction model⁴ for stratification of women within risk categories.

Data collected for PERSPECTIVE hold great value for future breast cancer research, including clinical data that characterize each breast cancer case associated with sequenced data. Secondary uses of data from PERSPECTIVE for as-yet unspecified research promises to enhance our knowledge of breast cancer etiologies without duplicating data-gathering efforts. The importance of making data available for secondary use in future research is now widely recognized among the scientific community and policy makers.^{5–7} Moreover, oncogenomic research constitutes one of the most promising fields for secondary use of data. The UK government 100,000 Genomes Project to sequence 100,000 genomes with a focus on cancer and rare disease⁸ will allow such access to data for further research.

Sharing research data with the scientific community can be mandatory. Indeed, data sharing is a requirement of the funders of PERSPECTIVE. It is both a scientific and an ethical responsibility. In 2015, efficient data sharing with a long-term vision – enabling the secondary use of data for years to come – should be considered standard practice for large-scale genomic research projects.

In this paper, we first illustrate how disseminating data is a core principle of PERSPECTIVE, and second, explain the ethical aspects of our data-sharing framework. Finally, we discuss a few ethical challenges we encountered in implementing this framework, and suggest how ongoing innovations might help to facilitate data sharing in future genomic research projects. Ethical challenges discussed include: 1) management of ethics approvals across jurisdictions, 2) return of results to research participants, and 3) research incentives and recognition for data production.

Data dissemination as a core principle

Storing and sharing data that are collected and produced in the context of research maximizes their utility, given the possibility of long-term data storage and multiple reuse. Providing access to these data is a return on the investment that society expects in exchange for its support of the research enterprise both financially and ideologically. Access to research data is also required to ensure that the full benefits of altruistic sample donation from research participants are realized, while allowing secondary uses of research data can make “actionable” the ethical principle of maximization of benefits.

The 2014 “International Charter of principles for sharing bio-specimens and data”⁹ reflects the beginning of an international consensus supporting global data-sharing practices. Authors from several countries (Sweden, Italy, Canada, USA, Australia, France, UK, and Germany) contributed to the writing of the charter, which is the result of negotiations with many stakeholders, including patient representatives, legal experts, ethical experts, and industry representatives. The Charter recognizes that data are maximized when they are shared, thereby accelerating research and yielding benefits to current and future patients. In the same year, the Global Alliance for Genomics and Health (GA4GH) adopted the “Framework for responsible sharing of genomic and health-related data”¹⁰ that provides guidance for the responsible sharing of data. Involving nearly 300 organizational members from 30 countries (eg, Japan, Brazil, China, Finland, Turkey, Mexico, India, Europe, USA, Canada, Italy, etc), the GA4GH aims to catalyze efforts that accelerate the potential of genomic medicine to advance human health.¹¹ The Framework acknowledges that sharing genomic and health-related data is key to ensuring continued progress in our understanding of human health and disease and is enshrined in the GA4GH constitution. Most importantly, it is built on two legally recognized human rights: the right of citizens to benefit from advances in science and of scientists to be recognized for their work. Finally, the *OECD Principles and Guidelines for Access to Research Data from Public Funding*,¹² published in 2007 by the Organization for Economic Co-operation and Development (OECD), also testify to a broader international movement in support of data sharing. OECD includes countries from North and South America, Europe, and Asia-Pacific, with also a few emerging countries like Mexico, Chile, and Turkey. Because data sharing can increase the returns of public investment in research, the OECD guidelines advocate for data sharing and open

access to publicly funded projects. The guidelines promote “openness and the free exchange of ideas, information and knowledge”. The *Fort Lauderdale Statement*¹³ adopted in 2003 and the Bermuda Principles¹⁴ of 1996 are other earlier examples of this evolution through the rapid prepublication release of research data, especially for DNA sequences.

Indeed, PERSPECTIVE was developed in the spirit of this movement. It complied with the funding agency policy to share data “in a timely fashion with minimal or no restrictions”¹⁵ through developing a concerted data-sharing plan concurrently with the core research proposal. PERSPECTIVE’s data-sharing plan embraced the principle of open science, releasing data with minimal restrictions.^{16,17}

Data sharing ethical framework

An ethical framework for data sharing involves decisions on how, when, to whom, why, etc, data should be disseminated. It precedes choices about technology or platforms used to implement data workflow and analysis, the latter benefiting from many current available options (eg, Galaxy,¹⁸ WebLab).¹⁹ Making data sharing effective in a large-scale international project requires a framework to manage related issues. Privacy and security protection are crucial. Despite its desirability, sharing within the scientific community can be a considerable ethical challenge. Although publicly available data may be considered private to an individual (eg, voters list, court decisions, salary and the expenses of elected representatives), laws govern the “publicity” of such data. The same may not hold for research. Certainly, genomic data derived from research participant samples may reveal personal information that requires a strict level of privacy protection. Genomic data may also provide information regarding actual or future health status. This concern is particularly acute in PERSPECTIVE where genomic data are associated with a wide range of phenotypic data (eg, diagnosis, age, ethnicity, location, disease in family members, etc) when combined, could possibly identify an individual. Data sharing is further impeded when participants donate samples with express limitations on their future use, whether geographical or commercial, or restrictions on the types of research use.

The ethical framework for data sharing in PERSPECTIVE was built on an existing research platform, the Breast Cancer Association Consortium (BCAC).³ Bringing together nearly 90 research units involved in breast cancer research, BCAC has been central to many discoveries related to breast cancer susceptibility.^{20–22} It uses a central database located in the Genetic Epidemiology Unit at the Department of Public

Health and Primary Care of the University of Cambridge (Cambridge, UK). To a great extent, BCAC data are used in PERSPECTIVE, or are derived from the samples the consortium provides. At the end of each phase of PERSPECTIVE, the data used and produced will be stored in the BCAC central database. This involves not only raw data, but also data generated from biological analysis and computational analysis: clinical data, phenotypic data, genotyping data, sequencing data, etc. Whether these data are usually called “raw data”, “clinical data”, or “results” is not relevant either for management of access or for the application of required privacy protection mechanisms. Considered on a continuum, all these data are the product of previous work, such as interviews with a patient, analysis of a mammogram, sample sequencing, etc.

Residual samples, if any, will be returned to their collection site (samples are not centrally stored or destroyed). Based on a decade’s worth of research collaboration, BCAC has a long tradition of developing infrastructures, processes, and guidelines for facilitating data management and sharing. It was therefore an obvious choice to build upon their existing data-sharing platform.

In conformity with the BCAC guidelines, a researcher wishing to access data should propose a specific type of analysis (eg, particular single-nucleotide polymorphism [SNP], pathway, or phenotype). Then, the Data Access Committee examines the proposal and decides on whether to provide an access authorization (see Figure 1). Composed of the three BCAC working group chairs (ie, Survival and Pathology, Mammographic Density, and Risk Factors), the main grant holders of BCAC, plus up to three additional BCAC investigators on a rotational basis, this committee also helps to identify overlaps, suggest potential merges, and organize proposals. If the Data Access Committee grants access, the researcher obtains the right to lead the analyses and publish as lead author. Both BCAC members and nonmembers can apply to access the data. Nonmembers, however, are subjected to a delay of upwards of 12 months after the completion of genotyping before they may publish results. This 1-year delay provides an equitable recognition for those who devoted time and energy to data collection and analysis to exercise their rights to publish.

Although workable for BCAC, these current data-sharing guidelines were not entirely suitable for the PERSPECTIVE project. Adaptations were required to fully address the ethical issues specific to our project.

The most important issue was that of potential limitations on future uses of the data produced. Since PERSPECTIVE

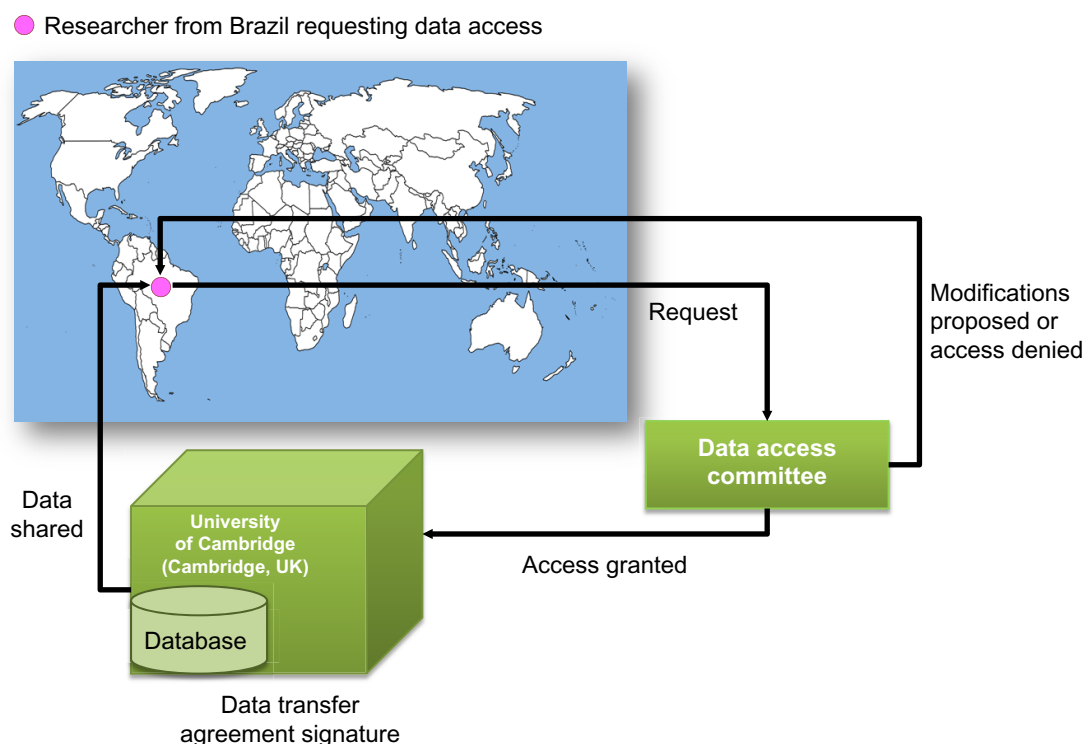


Figure 1 Disseminating the data in PERSPECTIVE process.

Abbreviation: PERSPECTIVE, Personalized Risk Stratification for Prevention and Early Detection of Breast Cancer.

only uses secondary data collected originally for a previous study, the data produced by PERSPECTIVE – and subsequently stored in the database – may be considered a “third” use. Since samples and data originate from 90 different studies around the world, the least we can expect is a variety of limitations on future use. Regulations – legal and ethical – regarding the use of data differ across jurisdictions. Limitations might arise from the consent obtained at the time of sample and data collection (sometimes decades ago). These limitations could relate to geographical areas (no foreign transfer); researcher affiliation (only those involved in the collection); status of the funding agencies (eg, public versus charity organization); aim of the research (eg, disease-specific); and length of storage (number of years mentioned). We want to both respect these limitations and also streamline the process so that only one data-sharing platform is used.

To address this issue, a contractual tool to regulate data sharing was developed. This tool comprises a set of research agreements that established the broad terms according to which data would be stored in the BCAC database and subsequently transferred for future research projects upon approved access. This Data Transfer Agreement template is mandatory for any transfer from the BCAC database. Existing BCAC data-sharing rules were also fully integrated as well as creating a customized clause for the inclusion of any

specific limitations. Therefore, each group that contributes data to PERSPECTIVE may specify its own limitations in the agreement, which all future researchers wishing to use these data must respect. Each data-access request must comply with any applicable limitations. Data obtained under different conditions at different times might be stored and shared using the same ethical framework, making easier a single access point for genomic cancer researchers.

The particular context of PERSPECTIVE explains why a public repository was not used. Firstly, this 5-year project required a sharing system that could both 1) store and disseminate the data within the team between each interrelated phase of the project and 2) progressively disseminate data to the scientific community (even before the end of 5 years). Secondly, the amount of clinical, phenotypic, and familial data collected from each individual would considerably challenge privacy in the context of an open-access repository. We could not ensure that anonymization would be sufficient to protect privacy considering specifically the amount of data collected and future use of family history for research. Finally, a myriad of intrinsic limitations to future use associated with data or samples at the time of collection (eg, disease-specific) could be difficult – or impossible – to manage by existing public repositories. Although a public repository was not an option suitable for PERSPECTIVE,

it remains a valuable option for many other projects. For instance, part of the data analyzed in collaboration with one phase of PERSPECTIVE will be put in dbGaP, as well as being in the BCAC database.

Actual ethical challenges and ongoing innovations

Like many projects attempting to expand their data-sharing practices, ethical approvals, return of results, and research incentives were the most pressing ethical issues encountered in the development of PERSPECTIVE's data-sharing framework. Below, we discuss from our experience how these ethical issues may affect a research protocol, and then how ongoing innovations can help manage these challenges.

Ethics approvals

Undue delays and wasted resources are not uncommon in the absence of harmonization between ethics review boards, both in the review process and in the application of ethical principles. Ethics review boards located outside one's regulatory home fall victim to similar delays as those located within the same jurisdictional borders. Obtaining multiple ethics approvals is common to most researchers leading multi-institutional projects, and many have expressed their concerns about how it affects the development of their project.^{23–27} Such evidence demonstrates the immediacy of considering solutions of, for example, a single, federated ethics review,²⁸ in contrast to the current system defined as “fragmented, inconsistent and inefficient”.²⁹

PERSPECTIVE also faced the challenge of multiple ethics approvals with samples and data provided from 90 studies and analysis conducted in several laboratories located in different countries. Creativity and strategy were crucial. Among the management strategies used, the establishment of adapted arrangements with the ethics review board of the lead investigator was one of the most helpful. We first divided the project into seven smaller substudies that regrouped similar issues. These numerous but simpler substudies were reviewed one after the other by the ethics review board. This arrangement avoided overburdening the process with too many complex issues at the same time. We also managed with that principal ethics review board to “recognize” foreign ethics approvals from many countries by providing the ethics review board confirmation about the status and the management of data, privacy protection, and the return of individual results in those countries. Mutual confidence between the ethics reviews boards involved and with the researchers greatly factored into our achievement.

It should be noted, however, that these negotiations are extremely time-consuming and involved lawyers from the institutions and not just the researchers.

Despite a uniform set of ethical guidelines applicable throughout Canada,³⁰ a national ethics review board does not yet exist. Some provinces have put in place initiatives to harmonize local ethics approval, but this is often restricted to a particular research field. For instance, Ontario has established a central ethics review board for all clinical trials in oncology.³¹ But nonclinical trial research has not benefitted from such centralization. For example, in 2015, Québec launched its new harmonization framework to simplify the multisite ethics approval process.³² Now, an ethics approval issued from a Québec health center will be automatically recognized by any other Québec health center. But, once again, harmonization in Québec will not be universal, as some ethics review boards (eg, universities and pharmaceutical companies) wished to remain excluded from this reciprocal recognition system.

More promising for large-scale international genomic research is the initiative led by the GA4GH. Its Regulatory and Ethics Working Group is currently working on the development of an ethics review equivalency concept to allow for mutual recognition of ethics reviews. The proposal is to foster ethics review equivalency to harmonize ethics review policies, foster ethical conduct, and ensure compliance for data-driven research projects.^{33,34} This “equivalency” concept may include a central ethical review board that could operate on an international basis with national input and response. This ambitious and innovative project could offer an alternative solution since inefficiency, cost, inconsistency, and redundancy are currently the norm in ethics approvals of genomic research for international consortia. The development, adoption, and implementation of this concept or, perhaps, a more federated model will require the involvement of many countries, funders, universities, and researchers.

Return of results

As a project involving genomic analyses, including whole-exome sequencing, PERSPECTIVE must confront the issue of return of individual results. Whether individualized results produced under the auspices of research should be communicated to individual research participants is one of the more controversial questions in contemporary bioethics, and is considered “one of the most difficult challenges facing investigators.”³⁵ The debate is considered a key frontier in research and practice because it is a question of translational science.³⁵

The American College of Medical Genetics and Genomics (ACMG) has weighed in on the debate, releasing its recommendations regarding when incidental findings in whole-exome and whole-genome sequencing should be returned.³⁶ Their mandatory return-of-results policy based on a panel of 57 genetic conditions raised strong opposition.^{37,38} ACMG then published a clarification³⁹ and finally revised its guidelines to offer an opt-out option of receiving incidental findings.⁴⁰ Although these recommendations are intended for the clinical context, they impact research. Indeed, the field of oncology demonstrates the ever-narrowing gap between research and care through translational initiatives.⁴¹ The frontier between research and clinic is diminishing in the push toward big data.⁴²

The context is very different in research where legacy samples are used. In PERSPECTIVE, samples and data were previously collected at different times, often with different conditions and objectives. Factors influencing participants' initial intent to participate in research also affect the decision to return individual results depending on the context. Additionally, institutional factors can dictate procedures for disclosure as well, including whether and if the possibility of incidental findings was discussed with participants at the time of consent, the clinical validity and utility of the result, and the disclosure policies at the collection site.⁴³ Since it is possible that data collection occurred in different contexts and time periods, we agreed to apply a case-by-case approach. Each sample provider complies with local ethics and legal guidelines, in accordance with the consent expressed by the participant. Thus, in the same project, some participants were entitled to receive some breast cancer-related results and others not based on their preference and existing policies. Of course, researchers can contact and re-consent where possible and practical, or they can get a waiver of re-consent from an ethics committee if certain situations are met.

It could be hazardous to propose a uniform standard for the return of results obtained from secondary (retrospective) use of data and samples in an international context. We can safely affirm, however, that each situation should be evaluated on an individual basis to ensure that return (or not) conforms to ethical and legal policies and the particular context of the research. But the use of data and samples to be collected in the future (prospective use) could benefit from current innovations in the ethics framework for return of results. In Canada, researchers have an ethical obligation to disclose any material incidental findings discovered in the course of the research.³⁰ Incidental findings are considered "material" if they have significant welfare implications for the well-being

of a participant. Researchers need to have their plan to return results or not approved by an ethics committee.

Prospective projects can hope to benefit from lessons learned from the past. New projects can turn to emerging ethical standards for data sharing, including the proactive management of incidental findings. The recently published "International Charter of principles for sharing bio-specimens and data" provides principles in collaboration with relevant stakeholders to do so: "participants should be allowed to have some options to express choices, or at least be provided clear information on the policy for return of incidental findings".⁹

Research incentives and recognition

One of the emerging ethical issues related to the use of previously collected data and samples, or biorepositories, is how to properly acknowledge those involved in different phases of the research. To successfully achieve data sharing, novel means must be developed to appropriately recognize and incentivize data generators.⁴⁴ Many have underlined this fundamental issue related to rewarding researchers in the recent context of data sharing at large scales, such as the UK Royal Society, which called for a fundamental shift in the scientific culture,⁴⁵ or the UK Expert Advisory Group on Data Access, who concluded that research culture does not provide adequate rewards for researchers.⁴⁶ It bears repeating that the GA4GH has made the dual human rights of citizens to benefit from scientific advances and of scientists to be recognized, the foundation for its framework for responsible sharing of genomic and health-related data.¹⁰

Advancement in research is closely related to research incentives and recognition. On the one hand, there is the traditional way of providing researcher incentives: intellectual property rights managed according to legal frameworks under specific jurisdictions. Patents, licenses, and trade secrets are part of this system for the recognition of researchers' work and contributions. They have been associated with genetic discoveries for decades. However, whether this system serves society's best interest is subject to vigorous debate.^{47,48} On the other hand, there are many challenges to this formal system and whether it ensures fair recognition of researchers involved in infrastructure science such as databases or biorepositories.^{45,46}

This issue of due recognition for researcher involvement was discussed during the development of PERSPECTIVE. As previously explained, PERSPECTIVE is built on the principle of open science, therefore data and results should be released with minimal restrictions.¹⁵ Before publishing

any results from PERSPECTIVE, the PERSPECTIVE Steering Committee shall determine whether there is an opportunity to patent or otherwise commercialize the content in respect of the applicable data and resource policies. These policies include, for instance, the possibility to use intellectual property rights defensively or strategically to better promote and disseminate the findings of the project to the global community. The current patent framework was not considered well suited for the type of collaboration planned in PERSPECTIVE. Managing patent ownership would have been a very complex issue considering involvement of hundreds of researchers and the sharing of samples and data from more than 90 repositories.

Expectations of endless and unproductive legal fights on patent issues related to genomics, such as those related to the Myriad Genetics BRCA gene patent,⁴⁹ the Long QT syndrome genetic test patent,⁵⁰ and the Canavan disease gene patent,⁵¹ prompted the PERSPECTIVE team to adopt the open science approach. Monopolies on clinical genetic technology can directly affect the quality of patient care.⁴⁷ Business-oriented strategies and monopolies can undermine the capacity of physicians to estimate patient risk.⁵² The cost of commercial patented genetic tests could be an important barrier for people in low-income countries, such as the case for noninvasive prenatal genetic testing.⁵³ Wide dissemination of PERSPECTIVE results – with the potential to enhance health care and stimulate more research – obtained through public funding, was judged to be in accordance with the ethical principles of maximization of benefits and reciprocity.

PERSPECTIVE will collect an important amount of raw data (eg, genotypic and phenotypic information) and will produce a large amount of raw data by sample genotyping and sequencing. However, considered alone, raw data or genomic data usually cannot be protected by patent or copyright. Therefore, the challenge is how to reward the producers of valuable datasets, as PERSPECTIVE contributors, whose data become the foundation for new analyses and understandings that researchers propose.⁵⁴ Bioresources are built from devoted scientists that used their leadership to get together expertise, innovative ideas, collaboration, and financial resources. But career metrics and assessment of productivity are still based on publications.⁴⁴ Creation of databases for use by others is not yet an achievement taken into account by funders and universities. On the one hand, broad data and sample sharing is a scientific imperative. On the other hand, it would be naïve to ignore the intellectual capacity and resources necessary to produce and make that

data available for researcher use. To the latter point, scientists deserve to have their contribution properly recognized, including when repositories are used.^{55,56} This expectation is consistent with the principles of a human rights approach⁵⁷ to research and funding rules for health research (eg, research funding is based on work previously published).⁵⁸ Thus, equitable recognition of the involvement of a researcher is crucial to monitoring both incentives and sustaining scientific innovation.

To address the issue of researcher recognition in PERSPECTIVE, the set of guidelines developed for BCAC were applied. Among these guidelines, some help manage researcher recognition and support research incentives:

- grant first option to researchers who propose to analyze a specific analysis to lead analyses and to publish results with lead authorship;
- give priority to BCAC members to publish results for a 12 month period after the completion of genotyping (or earlier, with special agreement);
- recognize the contribution of researchers who make specific data available by offering collaborations in carrying out the analyses or in writing the scientific paper, and/or granting authorship; and,
- recognize through authorship the involvement of individuals that provide data/samples, lead authors, members of the writing group or other working groups, and collaborators. In all cases, paper authorship must respect the recommendations of the International Committee of Medical Journal Editors that is, authorship should be based on substantial contribution to data acquisition and analysis.⁵⁹

BCAC guidelines are an illustration, hopefully suitable for the context of breast cancer genomics research, and the recognition of sample and data collection could be integrated into general research incentivization structures.

More dialogue and ethical policy initiatives are needed to develop management policy fully adapted to biorepositories and research databases. In all cases, a very delicate balance between data/samples release and researcher recognition should be maintained. Here too the “International Charter of principles for sharing bio-specimens and data”⁹ advocates that data and sample sharing should follow criteria for properly acknowledging researcher contributions through rules of authorship and intellectual property rights. Proposed rules are set depending on the source of data and samples, and reflect the level of engagement of people who contribute to data generation. For instance, while data from governmental administrative databases should not lead to

authorship or special recognition, requests for processed data such as genome-wide association study data could include an offer of coauthorship and a note in the acknowledgments section of future papers. The GA4GH Framework for responsible sharing of genomic and health-related data¹⁰ makes “Recognition and Attribution” one of the core elements of responsible data sharing. The Framework supports the development of data-sharing systems “which provide due credit and acknowledgement of all who contributed to the results”.¹⁰ It also recommends that recognition and attribution should be extended both to primary and secondary purposes, or downstream uses, and applications. The UK Royal Society has issued the *Science as an Open Enterprise* report on how the conduct of science needs to adapt to the new era of information technology.⁴⁵ Recognizing the reward system for researchers as an important issue to be addressed in order to promote data sharing, the recommendations propose to:

- recognize data communication as an important criterion for career progression and reward;
- reward open data on the same scale as journal articles and other publications;
- include measures for assessment that reward collaborative ways of working; and,
- improve the communication of research data by recognizing those who could maximize usability and good communication of the data.⁴⁵

The UK Expert Advisory Group on Data Access has identified areas in which incentives might be required to support researchers sharing data and to give due recognition to those who do so.⁴⁶ One of the recommendations of this advisory group is that research funders recognize the contribution of those who generate and share high-quality datasets, including as a formal criterion for assessing researchers’ achievements. This Group also suggests that contributions of both early-career researchers and data managers (the two most disadvantaged groups) be recognized and valued appropriately.

Conclusion: facilitating future data sharing

Scientists recently identified large consortia that combine data from different projects for joint analysis as a successful model of data sharing.⁶⁰ It is similar to PERSPECTIVE’s association with BCAC. Although consortia can provide tremendous support for efficient data sharing, they have some limits. One of the most obvious limits is at the same time a notable advantage: they are specialized. An

ethical framework adapted specifically to a narrow field of research, such as breast cancer, may not be applicable to another field of research, such as mental or communicable diseases. For instance, policy on the return of results to individuals applicable to breast cancer research may not be suitable elsewhere. Another challenge for consortia is the capacity to expand data sharing outside of the consortium: data sharing is considered achieved when the global research community can access the database. Therefore, efforts need to be devoted by consortia to enlarging their sharing capacity and to expanding their users to include the wider research community. Finally, major discrepancies between multiple, national ethical policies seriously challenge the capacity of consortia to efficiently achieve data sharing.

Disseminating data to the scientific community in large-scale international research projects presents a real ethical challenge. New approaches to consent, privacy protection, and intellectual property need to be developed to implement genomic technologies into personalized medicine.⁴⁷ In the absence of an effective ethical framework for rapidly sharing data across countries, researchers must be innovative. Human resources devoted to this “innovation” and its financial requirement are not frivolous. The lack of expertise and resources required to facilitate efficient data sharing remains a formidable barrier.

There are reasons to be optimistic, however, many of which are highlighted in this paper. Government agencies are increasingly prioritizing data sharing by making it an eligibility requirement to receive research funds, and innovative projects such as PERSPECTIVE are emerging which are founded on broad data-sharing practices and which encourage collaborative science. With enhanced data sharing comes the more pressing need to resolve issues of privacy protection, security, and ethics approvals not just within our regulatory homes, but beyond our jurisdictional borders. While retrospective projects may require more concerted harmonization efforts and tool development to facilitate data sharing, prospective projects may be better suited to benefit directly from the efforts hopefully exemplified by PERSPECTIVE and other similar data-sharing projects.

Acknowledgments

PERSPECTIVE is supported by the Québec Breast Cancer Foundation, by the Government of Canada through Genome Canada and the Canadian Institutes of Health Research, and by the Ministère de l’Économie, de l’Innovation et des

Exportations du Québec through Genome Québec. The authors would like to acknowledge Ms Shuang Shuang for her research assistance and Ms Vaso Rahimzadeh for the revision of the text and her helpful suggestions.

Disclosure

The authors report no conflicts of interest in this work.

References

- McNutt M. Data, eternal. *Science*. 2015;347(6217):7.
- Robinson PN. Genomic data sharing for translational research and diagnostics. *Genome Med*. 2014;6(9):78.
- Studies providing samples are members of the Breast Cancer Association Consortium (members list provided on the website). Available from: <http://apps.ccge.medschl.cam.ac.uk/consortia/bcac/index.html>. Accessed June 22, 2015.
- BOADICEA risk estimation on the World Wide Web. Cambridge: Centre for Cancer Genetics Epidemiology, Department of Public Health and Primary Care, University of Cambridge; 2013. Available from: <http://ccge.medschl.cam.ac.uk/boadicea/>. Accessed June 22, 2015.
- Poldrack RA, Gorgolewski KJ. Making big data open: data sharing in neuroimaging. *Nat Neurosci*. 2014;17(11):1510–1517.
- den Eynden VV, Corti L, Woollard M, Bishop L, Horton L. *Managing and Sharing Data: Best Practice for Researchers*. 3rd ed. Colchester: UK Data Archive, University of Essex; 2011.
- Walport M, Brest P. Sharing research data to improve public health. *Lancet*. 2011;377(9765):537–539.
- Genomics England. The 100,000 Genomes Project [webpage on the Internet]. London: Genomics England. Available from: <http://www.genomicsengland.co.uk/the-100000-genomes-project/>. Accessed February 9, 2015.
- Mascalzoni D, Dove ES, Rubinstein Y, et al. International Charter of principles for sharing bio-specimens and data. *Eur J Hum Genet*. 2015;23(6):721–728.
- Global Alliance for Genomics and Health. Framework for responsible sharing of genomic and health-related data. 2014. Available from: <http://genomicsandhealth.org/about-the-global-alliance/key-documents/framework-responsible-sharing-genomic-and-health-related-dat>. Accessed June 18, 2015.
- Global Alliance for Genomics and Health [homepage on the Internet]. Toronto: The Global Alliance for Genomics and Health. Available from: <http://genomicsandhealth.org/about-the-global-alliance/mission-founding-principles>. Accessed June 22, 2015.
- OECD Principles and Guidelines for Access to Research Data from Public Funding. Organisation for Economic Co-Operation and Development; 2007. Available from: <http://www.oecd.org/science/sci-tech/oecdprinciplesandguidelinesforaccesstoresearchdatafrompublicfunding.htm>. Accessed February 9, 2015.
- Sharing Data from Large-Scale Biological Research Projects: A System of Tripartite Responsibility. Fort Lauderdale: The Wellcome Trust; 2003. Available from: <http://www.genome.gov/pages/research/wellcomereport0303.pdf>. Accessed February 9, 2015.
- First International Strategy Meeting on Human Genome Sequencing. Summary of Principles Agreed at the First International Strategy Meeting on Human Genome Sequencing (reported by the Human Genome Organization), Southampton (Bermuda), 25–28 February 1996; Available from: http://www.casimir.org.uk/storyfiles/64.0.summary_of_bermuda_principles.pdf. Accessed June 18, 2015.
- Data Release and Resource Sharing. Genome Canada; 2008.
- Destro Bisol G, Anagnostou P, Capocasa M, et al. Perspectives on Open Science and scientific data sharing: and interdisciplinary workshop. *J Anthropol Sci*. 2014;92:179–200.
- Giglia E, Swan A. Open Access to data for a new, open science. *Eur J Phys Rehabil Med*. 2012;48(4):713–716.
- Galaxy [homepage on the Internet]. Available from: <http://galaxyproject.org/>. Accessed April 20, 2015.
- WebLab [homepage on the Internet]. Beijing: The WebLab Team. Available from: <http://weblab.cbi.pku.edu.cn/>. Accessed April 20, 2015.
- Breast Cancer Association Consortium. Commonly studied single-nucleotide polymorphisms and breast cancer: results from the Breast Cancer Association Consortium. *J Natl Cancer Inst*. 2006;98(19):1382–1396.
- Easton DF, Pooley KA, Dunning AM, et al. Genome-wide association study identifies novel breast cancer susceptibility loci. *Nature*. 2007;447(7148):1087–1093.
- Michailidou K, Hall P, Gonzalez-Neira A, et al. Large-scale genotyping identifies 41 new loci associated with breast cancer risk. *Nat Genet*. 2013;45(4):353–361.
- Hébert P, Saginur R. Research ethics review: do it once and do it well. *CMAJ*. 2009;180(6):597–598. English, French.
- Gold JL, Dewa CS. Institutional review boards and multisite studies in health services research: is there a better way? *Health Serv Res*. 2005;40(1):291–307.
- Silverman H, Hull SC, Sugarman J. Variability among institutional review boards' decisions within the context of a multicenter trial. *Crit Care Med*. 2001;29(2):235–241.
- Maxwell DJ, Kaye KI. Multicentre research: negotiating the ethics approval obstacle course. *Med J Aust*. 2004;181(8):460–461.
- Ezzat H, Ross S, von Dadelszen PV, Morris T, Liston R, Magee LA; CPN Collaborative Group. Ethics review as a component of institutional approval for a multicentre continuous quality improvement project: the investigator's perspective. *BMC Health Serv Res*. 2010;10:223.
- Whiteman DC, Webb PM, Purdie DM, Green AC. National ethics committee urgently needed. *Med J Aust*. 2003;178(4):187.
- Knoppers BM. International ethics harmonization and the global alliance for genomics and health. *Genome Med*. 2014;6(2):13.
- Canadian Institutes of Health Research, Natural Sciences and Engineering Research Council of Canada, Social Sciences and Humanities Research Council of Canada. *Tri-Council Policy Statement: Ethical Conduct for Research Involving Humans*. 2nd ed. Ottawa: Government of Canada; 2014.
- Ontario Cancer Research Ethics Board [webpage on the Internet]. Toronto: Ontario Institute for Cancer Research (OICR). Available from: <http://oicr.on.ca/oicr-programs-and-platforms/ontario-cancer-research-ethics-board>. Accessed February 10, 2015.
- Ministère de la Santé et des Services sociaux. Framework for Public Health and Social Services Institutions to authorize research conducted at more than one site. Available from: http://www.mcgill.ca/medresearch/files/medresearch/msss_framework_-_english_translation_2014-12-19.pdf. Accessed June 18, 2015.
- Dove ES, Knoppers BM, Zawati MH. Towards an ethics safe harbor for global biomedical research. *J Law Biosci*. 2014;1(1):3–51.
- Dove ES, Knoppers BM, Zawati MH. An ethics safe harbor for international genomics research? *Genome Med*. 2013;5(11):99.
- Wolf SM. Return of individual research results and incidental findings: facing the challenges of translational science. *Annu Rev Genomics Hum Genet*. 2013;14:557–577.
- Green RC, Berg JS, Grody WW, et al; American College of Medical Genetics and Genomics. ACMG recommendations for reporting of incidental findings in clinical exome and genome sequencing. *Genet Med*. 2013;15(7):565–574.
- Burke W, Antommarchia AH, Bennett R, et al. Recommendations for returning genomic incidental findings? We need to talk! *Genet Med*. 2013;15(11):854–859.
- Ross LF, Rothstein MA, Clayton EW. Premature guidance about whole-genome sequencing. *Per Med*. 2013;10(6).
- American College of Medical Genetics and Genomics. Incidental findings in clinical genomics: a clarification. *Genet Med*. 2013;15:664–666.
- ACMG Board of Directors. ACMG policy statement: updated recommendations regarding analysis and reporting of secondary findings in clinical genome-scale sequencing. *Genet Med*. 2015;17:68–69.

41. Doroshow JH, Kummar S. Translational research in oncology – 10 years of progress and future prospects. *Nat Rev Clin Oncol*. 2014;11:649–662.
42. Chute CG, Ullman-Cullere M, Wood GM, Lin SM, He M, Pathak J. Some experiences and opportunities for big data in translational research. *Genet Med*. 2013;15:802–809.
43. Rahimzadeh V, Avar D, Sénécal K, Knoppers BM, Sinnett D. To disclose, or not to disclose? Context matters. *Eur J Hum Genet*. 2015;23(3):279–284.
44. Kaye J, Hawkins N. Data sharing policy design for consortia: challenges for sustainability. *Genome Med*. 2014;6:4.
45. *Science as an Open Enterprise*. London: The Royal Society; 2012. Available from: <https://royalsociety.org/~media/policy/projects/sape/2012-06-20-saoe.pdf>. Accessed April 24, 2015.
46. *Establishing Incentives and Changing Cultures to Support Data Access*. Expert Advisory Group on Data Access (EAGDA); 2014. Available from: http://www.wellcome.ac.uk/stellent/groups/corporatesite/@msh_peda/documents/web_document/wtp056495.pdf. Accessed February 10, 2015.
47. Ormond KE, Cho MK. Translating personalized medicine using new genetic technologies in clinical practice: the ethical issues. *Per Med*. 2014;11(2):211–222.
48. Sayres LC, Allyse M, Goodspeed TA, Cho MK. In the public Interest? *Sci Transl Med*. 2012;4(144):144fs23.
49. So D, Joly Y. Commercial opportunities and ethical pitfalls in personalized medicine: a myriad of reasons to revisit the Myriad Genetics saga. *Curr Pharmacogenomics Person Med*. 2013;11(2):98–109.
50. Faculty of Law, McGill University. Press release: Children's Hospital Sues over Gene Patent, with Help from top McGill University Patent Law expert [press release]. Montréal: McGill University; 2014 [November 3]. Available from: https://www.mcgill.ca/law/files/law/gold-richard-cheo-mcgill_release-03-11-2014-eng.pdf. Accessed May 7, 2015.
51. Colaiaanni A, Chandrasekharan S, Cook-Deegan R. Impact of gene patents and licensing practices on access to genetic testing and carrier screening for Tay-Sachs and Canavan disease. *Genet Med*. 2010;12(4 Suppl): S5–S14.
52. Conley JM, Cook-Deegan R, Lázaro-Muñoz G. Myriad after Myriad: the proprietary data dilemma. *N C J Law Technol*. 2014;15(4):597–637.
53. Allyse M, Minear MA, Berson E, et al. Non-invasive prenatal testing: a review of international implementation and challenges. *Int J Womens Health*. 2015;7:113–126.
54. Kaye J, Heeney C, Hawkins N, de Vries J, Boddington P. Data sharing in genomics – re-shaping scientific practice. *Nat Rev Genet*. 2009;10:331–335.
55. Bravo E, Calzolari A, De Castro P, et al. Developing a guideline to standardize the citation of bioresources in journal articles (CoBRA). *BMC Med*. 2015;13:33.
56. The Bioresource Research Impact Factor (BRIF). Available from: <http://www.gen2phen.org/groups/brif-bio-resource-impact-factor>. Accessed February 11, 2015.
57. Knoppers BM, Harris JR, Budin-Ljøsne I, Dove ES. A human rights approach to an international code of conduct for genomic and clinical data sharing. *Hum Genet*. 2014;133(7):895–903.
58. Gitter DM. The challenges of achieving open-source sharing of biobank data. *Biotechnol Law Rep*. 2010;29:623–635.
59. *Recommendations for the Conduct, Reporting, Editing, and Publication of Scholarly Work in Medical Journals*. International Committee of Medical Journal Editors (ICMJE); 2014. Available from: <http://www.icmje.org/icmje-recommendations.pdf>. Accessed February 10, 2015.
60. Simpson CL, Goldenberg AJ, Culverhouse R, et al. Practical barriers and ethical challenges in genetic data sharing. *Int J Environ Res Public Health*. 2014;11(8):8383–8398.

Advances in Genomics and Genetics

Publish your work in this journal

Advances in Genomics and Genetics is an international, peer reviewed, open access journal that focuses on new developments in characterizing the human and animal genome and specific gene expressions in health and disease. Particular emphasis will be given to those studies that elucidate genes, biomarkers and targets in the development of new or improved therapeutic

interventions. The journal is characterized by the rapid reporting of reviews, original research, methodologies, technologies and analytics in this subject area. The manuscript management system is completely online and includes a very quick and fair peer-review system. Visit <http://www.dovepress.com/testimonials.php> to read real quotes from published authors.

Submit your manuscript here: <http://www.dovepress.com/advances-in-genomics-and-gene-expression-journal>

Dovepress